

**Estimating global ocean heat content changes in  
the upper 1800 m since 1950 and the influence of  
climatology choice\***

John M. Lyman<sup>1</sup>

Joint Institute for Marine and Atmospheric Research, University of  
Hawaii at Manoa, Honolulu, Hawaii, and NOAA/Pacific Marine  
Environmental Laboratory, Seattle, Washington

Gregory C. Johnson

NOAA/Pacific Marine Environmental Laboratory, Seattle, Washington

15 November 2013

Accepted for publication in *Journal of Climate*

---

\* Pacific Marine Environmental Laboratory Contribution Number 3884 and Joint Institute for Marine and Atmospheric Research Contribution Number 12-380.

<sup>1</sup> *Corresponding author address:* John Lyman, NOAA/Pacific Marine Environmental Laboratory, 7600 Sand Point Way N.E., Bldg. 3, Seattle, Washington 98115-6349, U.S.A.  
E-mail: John.Lyman@noaa.gov

## Abstract

Ocean heat content anomalies are analyzed from 1950 to 2011 in five distinct depth layers (0–100, 100–300, 300–700, 700–900, and 900–1800 m). These layers correspond to historic increases in common maximum sampling depths of ocean temperature measurements with time, as different instruments—mechanical bathythermograph (MBT), shallow expendable bathythermograph (XBT), deep XBT, early sometimes shallower Argo profiling floats, and recent Argo floats capable of worldwide sampling to 2000 m—have come into widespread use. This vertical separation of maps allows computation of annual ocean heat content anomalies and their sampling uncertainties back to 1950 while taking account of in situ sampling advances and changing sampling patterns. The 0–100-m layer is measured over 50% of the globe annually starting in 1956, the 100–300-m layer starting in 1967, the 300–700-m layer starting in 1983, and the deepest two layers considered here starting in 2003 and 2004, during the implementation of Argo. Furthermore, global ocean heat uptake estimates since 1950 depend strongly on assumptions made concerning changes in undersampled or unsampled ocean regions. If unsampled areas are assumed to have zero anomalies and included in the global integrals, the choice of climatological reference from which anomalies are estimated can strongly influence the global integral values and their trend. The sparser the sampling and the bigger the mean difference between climatological and actual values, the larger the influence.

## 39 1. Introduction

40 The world's oceans have absorbed roughly 90% of the anthropogenic warming from  
41 greenhouse gasses since the 1960s (Bindoff et al. 2007). This well-documented ocean  
42 warming occurs primarily in the upper 0–700 m (Domingues et al. 2008; Ishii and  
43 Kimoto 2009; Levitus et al. 2009), but warming has also been observed at intermediate  
44 and mid-depths, from 700 to 2000 m (von Schuckmann et al. 2009; Levitus et al. 2012),  
45 and in the abyssal ocean, well below 2000 m (Purkey and Johnson 2010; Kouketsu et al.  
46 2011).

47 Historically, ocean temperature sampling is inhomogeneous both by geographic  
48 region and depth from year to year, with major shifts in observing techniques and  
49 observational programs from decade to decade (Johnson and Wijffels 2011). Accurate but  
50 spatially sparse temperature measurements using reversing thermometers, sometimes to  
51 full depth, commenced after 1874. These thermometers were gradually replaced, starting  
52 in the 1960s, with conductivity-temperature-depth (CTD) instruments. Starting in the  
53 1930s, mechanical bathythermographs (MBTs) made it easier to measure the upper ocean  
54 temperature (sampling as deep as 145 to 300 m, more often shallower than deeper). They  
55 were widely used in the 1950s and 1960s. Starting in 1966, the expendable  
56 bathythermograph (XBT) began to replace the MBT for upper ocean thermal sampling,  
57 with shallow XBTs (sampling as deep as 460 m) dominant in the 1970s and 1980s, and  
58 deep XBTs (sampling as deep as 760 m) dominant in the 1990s. Like MBTs, XBTs are  
59 not very accurate (nominally  $\pm 0.1^{\circ}\text{C}$  for temperature and  $\pm 2\%$  for depth), but both  
60 comprise a large part of the observational record of upper ocean temperatures. The

61 profiling CTD floats of the Argo Program (Roemmich et al. 2009) began to provide  
62 accurate, year-round sampling of upper ocean temperature starting in 2000. The array  
63 started regionally, but became near-global and largely replaced the XBT for distributed  
64 ocean measurements by 2005. Early on, some of the floats only sampled to around 1000  
65 m, but as time has gone on, more floats have sampled as deep as the 2000-m Argo target.

66 For climate studies, ocean temperature data are used to estimate ocean heat  
67 content anomalies (OHCA). There are many sources of uncertainty in estimates of annual  
68 global integrals of OHCA. While reversing thermometers and CTDs are relatively  
69 accurate, XBTs have biases in both depth and temperature; different choices of XBT bias  
70 corrections lead to large uncertainties from at least 1993 to 2008 (Lyman et al. 2010;  
71 Palmer et al. 2010). Mechanical bathythermograph (MBT) bias corrections are equally  
72 important in earlier years (Gouretski and Reseghetti 2010; Ishii and Kimoto, 2009).  
73 Sampling uncertainty also contributes significantly to the overall uncertainty, and has  
74 strong dependence on time (Lyman and Johnson 2008; Domingues et al. 2008; Palmer  
75 and Brohan 2011) with larger uncertainties occurring for earlier years when the global  
76 ocean is not well sampled. Finally, uncertainties from mapping choices and climatology  
77 choice also contribute to the overall uncertainty. Mapping is variously done by bin  
78 averaging (Palmer et al. 2007; Gouretski et al. 2012), objective analysis (Ishii and  
79 Kimoto, 2009; Levitus et al. 2009), or reduced-space optimal interpolation (Domingues et  
80 al. 2008).

81 Climatologies representative of a single era or a long-term mean from which to  
82 estimate anomalies in ocean heat content have, historically, been difficult to compute  
83 because of sparse historical sampling. Only in recent years has the global ice-free ocean

all been sampled adequately in a given year or few years to construct a monthly climatology for a fixed time-period (e.g., Roemmich and Gilson 2009). Having a climatological baseline using data with a well-resolved seasonal cycle for a single, short time period (2005–2010), as is done here, can be very important in improving the accuracy of estimates of global integrals of ocean heat content anomalies in a historically sparsely sampled warming ocean, to avoid aliasing large signals from the seasonal cycle into the annual maps.

Here we estimate annual global integrals of ocean heat content in five distinct depth layers from 1950 through 2011. We focus on how sparse irregular sampling and choice of climatological baseline affects estimates of the global integral of OHCA. Here these anomalies are estimated relative to a monthly Argo-period climatology, with a well-resolved seasonal cycle collected over a relatively short time period. The five distinct depth layers (0–100, 100–300, 300–700, 700–900, and 900–1800 m) are chosen to capture the historical sampling patterns of MBTs, XBTs, and Argo floats. Since profiles often terminate at depths slightly shallower than the nominal maximum instrument sampling depth, layer bottom depths are chosen somewhat shallower than the nominal maximal sampling depths of MBTs (100 vs. 145–300 m), shallow XBTs (300 vs. 460 m), deep XBTs (700 vs. 760 m), and Argo floats (900 and 1800 vs. 1000 and 2000 m, respectively) to maximize the number of profiles spanning each layer yearly.

We further use sea surface height (SSH) as a proxy to estimate the OHCA sampling uncertainty in each of the depth intervals mentioned above, building on the work of Lyman and Johnson (2008) for a single 0–700-m layer. This study also builds on that study by making estimates of global OHCA values from in situ data for each layer as

outlined above. The earlier study used only proxy data to study sampling uncertainties, without producing actual estimates of OHCA from in situ data. Here we also quantify the effects of infilling on the global integral of OHCA by performing two integrations. One assumes zero anomalies in unsampled regions (the zero-infill mean, or ZIF) and the other that unsampled regions have anomalies equal to the average anomaly of sampled regions (the representative mean, or REP). We also build on previous work, presenting an analysis of how choices of baseline climatology can modify the effects of infilling assumptions on the global OHCA estimates.

## 2. Data

This study uses two sources of data. The Hadley Centre observational dataset EN3\_v2a (Ingleby and Huddleston 2007) is used, with XBT and MBT bias corrections of Ishii and Kimoto (2009) applied. All the profiling float data are removed from this dataset, leaving data from temperature profiles from bottle-station reversing thermometers, CTDs, XBTs, MBTs, and moorings. Profiling float data are obtained from the U.S. Argo Global Data Acquisition Center (GDAC), and data from floats with possible pressure drifts are removed (Barker et al. 2011). Delayed-mode data are used when available; otherwise real-time data are used. Only data with quality flags of “good” are retained.

We apply additional quality control when computing OHCA to ensure that each layer is adequately sampled. Profiles that do not extend to the bottom depth of a given layer are not considered for that layer. Furthermore, for data from a profile to be used in any given layer, its mean vertical measurement spacing must, on average, be closer than twice the spacing of NODC World Ocean Atlas standard levels from 0 to 100 m, and closer than 3 times the standard level spacing below 100 m.

### 3. Method

OHCA and its sampling uncertainty are computed following Lyman and Johnson (2008), except this time, the OHCA is computed from in situ data instead of proxy data. Aside from the quality control steps mentioned above and the use of in situ data, the other major differences are that OHCA is referenced to a seasonally varying monthly Argo climatology and both the OHCA and its sampling uncertainty are computed for five distinct depth layers (0–100, 100–300, 300–700, 700–900, and 900–1800 m) as detailed below. We discuss curves of the global integral of OHCA for 0–100, 0–300, 0–700, 0–900, and 0–1800-m layers. We compute these curves by adding the global integrals of OHCA for the appropriate sets of mapped layers.

#### *a. OHCA mapping*

Heat content in each of the five distinct layers is computed first for every profile. A new seasonal climatology of heat content with monthly resolution is computed using Argo data from 2005 to 2010. This climatology is removed from the heat content profiles to generate anomalies for each layer adequately sampled by the profiles. The results are then objectively mapped using a covariance function with both a large (~1000 km) scale and a small (~100 km) scale (Willis et al. 2004; Lyman and Johnson 2008), assuming a noise-to-signal variance ratio of 2.2. The mapping scales and noise-to-signal ratios could be refined to vary geographically in the future. Ocean temperature covariance scales vary with latitude, becoming elongated in the zonal direction near the equator (e.g., Roemmich and Gilson 2009). They may also vary with depth (e.g., Carton et al. 1996).

*b. Global integral of OHCA*

We examine two simple assumptions that can be made when computing a global integral of OHCA, again following Lyman and Johnson (2008). The first assumption is that anomalies are zero in regions that are not sampled, a common one for objective mapping. Integrals of the OHCA objective maps using this assumption have been referred to previously as simple integrals or means of the maps (Lyman and Johnson 2008). Here we will use the more descriptive term zero-infill mean (ZIF). This sort of assumption is commonly used when integrating objective maps (e.g., Ishii and Kimoto 2009; Levitus et al. 2012), and sometimes when summing bin-averaged data over the globe (Gouretski et al. 2012).

The second assumption is that the anomalies in regions that are not sampled are the average of anomalies in sampled regions, an assumption that has also been applied to global sums of bin-averaged data (Palmer et al. 2007). Integrals of OHCA maps using this assumption have been referred to as the weighted integral (Lyman and Johnson 2008). Here we will use the term representative mean (REP). Since objective maps also tend to relax back to zero in areas that are undersampled, making a representative mean using an objective map is slightly more involved than doing so using bin-averaged data. We apply the formalism of Lyman and Johnson (2008) to counteract the tendency of objective maps to relax to zero in data-poor regions when making a representative mean for an objective map. There are other more sophisticated methods for infilling gaps using satellite altimetry data either in an objective mapping formalism (e.g., Willis et al. 2004) or with reduced-space optimal interpolation (e.g., Domingues et al. 2008), but we do not address them here.



*c. Sampling uncertainty of global OHCA integrals*

Sampling uncertainties are computed in each mapped layer by using Aviso SSH as a proxy for OHCA (Lyman and Johnson 2008). Distinct spatially varying regression coefficients are computed between SSH and the OHCA content for each mapped layer using data from 1993 to 2011 in modified WMO squares (Willis et al. 2004). The regression coefficients are larger in the upper 0–100 m and decrease with depth, likely owing to decreasing vertical temperature stratification and eddy energy with depth.

When spatially integrating the SSH proxy of OHCA, areas that were not sampled are assumed to have the same anomalies as the area average anomaly of sampled regions. This representative mean yields an unbiased estimate of the global integral of Aviso SSH, bolstering confidence in the sampling uncertainties obtained by the method (Lyman and Johnson 2008).

SSH as a proxy for OHCA is most applicable in the layers spanning the thermocline, because SSH variability is dominated by the energetic temperature variability in these strongly stratified, mostly upper layers. The 0–300-m layer is highly energetic and explains ~92% of the SSH variability (Llovel et al. 2013). However, there still are strong correlations between SSH and temperature at 400 m (Fig 8. of Wijffels et al. 2008), with correlations  $> 0.50$  outside the tropics but lower values in the tropics where the thermocline is shallower than 400 m. For our choice of layers (Fig. 1) we find the highest global area-weighted correlation coefficient between SSH and OHCA occurs in the 100–300-m layer (0.54) and the lowest in the 900–1800-m layer (0.38). The low tropical correlations in the 300–700-m layer (below the tropical thermocline) are even

lower in the deepest 900–1800-m layer analyzed here (Fig. 1), but correlations also increase with increasing layer depth at some high latitude locations.

When adding sampling uncertainties of the distinct depth layers, uncertainties are assumed correlated in areas where both layers are sampled and assumed uncorrelated if sampling areas in different layers do not overlap. This procedure is detailed for a general linear mapping in Appendix A. These results are dependent on the spatial and temporal scales used in the mapping, as well as the ratio of the unresolved noise and signal variances.

When increasing length or time scales for covariance functions increases the amount of data used, the ratio of noise to signal variance will increase as well, so the uncertainty might not decrease. Furthermore, lengthening the temporal scale beyond annual cannot overcome aliasing of energetic modes of interannual ocean temperature variability, most prominently El Niño and La Niña, since these phenomena redistribute large quantities of ocean heat horizontally and vertically over time-scales shorter than a year (e.g. Johnson et al. 2000, Roemmich and Gilson 2011) and are linked to the seasonal cycle (e.g., Harrison and Larkin 1998, Larkin and Harrison 2001, Stuecker et al. 2013).

The sampling uncertainty for the REP is also a measure of how well that method fills in data gaps. The sampling uncertainty is computed by comparing sub-sampled OHCA SSH proxy curves with complete OHCA SSH proxy curves (Lyman and Johnson 2008). The complete SSH fields contain eddies, gyres, and some portion of the net warming signal in areas where there are no in situ data. The sampling uncertainty is a measure of how well the infilling represents missing variability that is contained in the complete SSH fields. The REP estimates produce a smaller sampling uncertainty than the

ZIF estimates because the former methodology incorporates a more conservative, but perhaps naïve, assumption about the contribution of unsampled regions than the latter (Lyman and Johnson 2008).

*d. Climatology baseline shifts*

We use a climatology based on Argo data from 2005 to 2010 (ClimArgo) because of Argo's even sampling throughout the year and its near-global, comparatively even spatial coverage of the upper 2000 m of the world's oceans. The resulting climatology is a modern one, and thus almost certainly warmer than a long-term mean over our analysis period or conditions in, say, 1950 (as far back in time as estimates are attempted here). Given the sparse sampling, how would the global integrals change if a different, colder climatology were used?

We investigate this question by introducing a second climatology, which is just the Argo climatology with values shifted uniformly by a global representative mean estimate of heat content anomaly for each distinct layer in 1950 (Clim1950). The general case, solved for any linear mapping and any mean offset in the climatology, is detailed in Appendix B. We choose a representative mean from 1950 so that the shift in the climatology spans a large but realistic range ( $m^{1950,l}$ , B3 evaluated at 1950), perhaps an upper bound for the modern historical record. This shifted Argo climatology is considerably colder than ClimArgo. The better to compare the resulting OHCA curves, the REP global average heat content anomaly from 1950 ( $m^{1950,l}$  A, B2 evaluated at 1950) is added back to the OHCA curves estimated using Clim1950. This addition shifts the REP curves estimated using Clim1950 to overlay the ClimArgo curves (B11). However,

shifts in the ZIF curves estimated using Clim1950 are dependent on the sampling in each year (B10) and do not overlay each other.

The uniform cooling shifts applied to each layer of ClimArgo to obtain Clim1950 only approximate the magnitude of the global ocean temperature changes from 1950 to 2005–2010. Actual global ocean temperature changes, even over several decades, exhibit strong regional variations (e.g., Gleckler et al. 2012, Levitus et al. 2012). Here we neglect the regional variations to concentrate on the simpler first order effects of uniform warming. By merely shifting the modern Argo climatology we also avoid issues that would arise from attempting to construct a global climatology during the sparsely sampled 1950s.

#### 4. Historical sampling

The historical record of vertical and horizontal structure of ocean temperature sampling is linked to the history of instrument development. The first worldwide survey of upper ocean temperatures, completed with min-max thermometers, was taken during the Challenger Expedition in 1872–1876. Those data have recently been compared to Argo float data from 2004–2010, finding statistically significant ocean warming over the intervening time interval (Roemmich et al. 2012). The development of the reversing thermometer in 1874 allowed measurement of temperature and depth over full ocean depths to a maximum accuracy of about 0.01°C and 0.5%, respectively (Warren 2008). However, reversing thermometers require a ship to stop, and historical measurements prior to the International Geophysical Year (1957–1958) are globally very sparse (Boyer et al. 2009).

The development of the MBT circa 1938 allowed measurements from a moving ship, although the ship had to slow to a few knots for best results, profiles were no deeper than 300 m, often only to 100 m or so, and temperature accuracy was probably not as good as the stated 0.1°C (Emery and Thomson 1998). MBT depth accuracy is at best 1% (Gouretski and Reseghetti 2010). Nonetheless, MBT profiles are the predominant upper ocean temperature observations from circa 1950 to 1969 (e.g., Johnson and Wijffels 2011). Their extensive use in the Northern Hemisphere in the Pacific and Atlantic oceans during that time period is evident in the average yearly 0–100-m layer sampling density (Fig. 2a). Bottle measurements that extend to depths of 1800 m in these same regions are far less extensive from 1950 to 1969 (Fig. 2e; i; m). The annual sampling fraction,  $f_t^y$ , (A2) of the 0–100-m layer exceeds 50% starting in 1956 (Fig. 3). The widespread use of the MBT allows that 50% sampling benchmark to be reached at this early date, but only for the uppermost, 0–100-m, layer.

From circa 1970 to 1989, shallow XBTs, profiling as deep as 460 m, are predominant contributors to the historical temperature record (e.g., Boyer et al. 2009; Johnson and Wijffels 2011). Accuracies of  $\pm 0.1^\circ\text{C}$  and  $\pm 2\%$  for temperature and depth, respectively, are specified for the XBT. However, systematic, time-dependent biases can approach or exceed these values (e.g., Gouretski and Reseghetti 2010). The use of the shallow XBT coincides with well-measured temperature (on an annual basis) in the northern hemispheres of all three oceans (now including the Indian Ocean) and extending to the 100–300-m layer (Fig. 2b; f). The Southern Hemisphere subtropics are also somewhat better, although still inadequately (in terms of annual averages) measured in these layers during this time. Reversing thermometer and CTD data collected during this

period increase temperature sampling in the 300–700, 700–900, and 900–1800-m layers (Fig. 2l; n). The annual sampling fraction of the 100–300-m layer exceeds 50% starting in 1967 (Fig. 3), as a result of the commencement of widespread use of the shallow XBT starting around that time.

Deeper XBTs (measuring frequently to at least 700 m, with measurement accuracies and biases similar to the shallow XBTs; e.g., Gouretski and Reseghetti 2010), while introduced earlier, become an important source of historical temperature data from circa 1990 to 2004 (e.g., Johnson and Wijffels 2011) during and after the World Ocean Circulation Experiment (WOCE). During this time period, adequate sampling extends into the Southern Hemisphere subtropics of all three oceans, even to the 300–700-m layer (Fig. 2c; g; k). The annual sampling fraction of the 300–700-m layer exceeds 50% starting in 1983 (Fig. 3) as a result of the increasingly widespread use of the deep XBT during the pre-WOCE large-scale measurement programs. This sampling coverage to 700 m peaks at a local temporal maximum of around 79% in 1994 and then declines to a local temporal maximum of 63% in 2001, perhaps because some XBT data are slow to filter into the data archival centers, but at least partly as a result of difficulties sustaining gains in ocean observations at the end of WOCE. Deep CTD stations occupied during WOCE further improve sampling in the 700–900 and 900–1800-m (Fig. 2o) layers (but still not adequately outside of the North Atlantic and western North Pacific oceans). These layers also show a similar decline between 1993 and 2000 with the end of WOCE (Fig. 3).

The Argo program of autonomous profiling CTD floats (Roemmich et al. 2009) began sampling the upper half of the ocean (nominally to 2000 m) in the year 2000. These floats have nominal temperature accuracy of 0.02°C, although usually much better,

and nominal pressure (hence depth) accuracy of 0.12%. By 2005 (and through 2010 and beyond) Argo achieved near-global coverage in the ice-free oceans excluding marginal seas down to 900 m (Fig. 2j). However, the very light surface waters of the tropics presented a challenge for some early Argo floats, and not all of them could profile there much beyond 1000 m, leaving that area slightly undersampled during 2005–2010 (Fig. 2p). As more capable floats that can profile from the surface to 2000 m even in the tropics replace the earlier, more limited floats, that coverage gap is gradually being filled. The annual sampling fraction of the 700–900-m layer exceeds 50% starting in 2003 (Fig. 3). That same milestone is reached for the 900–1800-m layer starting in 2004 as a result of the increasingly capable and widespread Argo array. By 2010, over 80% of the global ice-free oceans are sampled even to the 900–1800-m layer (Fig. 3).

Sampling of the ocean temperature below 2000 m (not treated here) is still very sparse, limited at present to shipboard CTDs. In most areas, comparisons of co-located oceanographic sections (e.g., Purkey and Johnson 2010) or data assimilation into models (e.g., Kouketsu et al. 2011) may be the most appropriate way to assess decadal changes below 2000 m, although it is possible to map multi-decadal changes in ocean heat content in the North Atlantic (e.g., Lozier et al. 2008; Mauritzen et al. 2012), historically the best-sampled ocean.

## 5. Climatologies and time-varying spatial coverage

The method used to integrate annual OHCA maps globally has a large impact on estimates of heat uptake since 1950 in the upper 0–1800 m of the global oceans. The ZIF estimates referenced to ClimArgo (Fig. 4, dashed grey lines) yield a smaller overall rate of heat uptake since 1950 than the REP estimates (Fig. 4, solid black lines) in all depth

ranges considered. The ZIF estimates referenced to the ClimArgo (presumably warm) climatology tend to relax back to zero when the data coverage for the integrated layers (Fig. 3) becomes sparse. This result occurs because the objective maps relax back to zero anomaly where data are sparse; zero anomaly in most locations will, presumably, tend to be on the warm side of the (poorly known) long-term average because the Argo climatology used is modern (2005–2010). In other words, the annual ZIF estimates using a modern climatology are pulled toward cooler values earlier in time by cooler temperature data, but for times and locations when and where data are sparse, they relax back to the mean, which is warm owing to the modern climatology used.

One way to demonstrate this effect is to use a cooler climatology, Clim1950, to produce contrasting curves for the ZIF estimates (Fig. 4, black dashed lines). Clim1950 is simply a shift of the Argo climatology to the colder mean 1950 values estimated using the representative mean anomalies for each layer. It may be about the coldest reasonable climatology that could be used for assessing heat content changes over this time period. The use of Clim1950 results in a much larger warming trend for the curves produced from ZIF estimates than does the use of ClimArgo. This difference arises solely because ZIF estimates assume zero anomalies in unsampled regions. With a climatology shifted to colder 1950 mean values, the large extents of the maps with sparse data in the early years all relax back to relatively colder 1950 mean values as a baseline, rather than the warmer Argo climatology values. In the more recent years, the larger sampled fraction of the global ocean (Fig. 3) means that ZIF estimates for either the colder Clim1950 or the warmer modern ClimArgo are more constrained by data to be similar (Fig. 4). Hence, for the ZIF estimates, the changing and sometimes sparse sampling patterns combine with



the climatology to produce trends that depend on the climatology values, especially if those trend estimates span times of sparse sampling. Shifting the Argo climatology using representative means between 1950 and 2000 produces similar shifts in the ZIF estimates (B10, plots not shown).

The size of the shifts in the ZIF estimates is dependent on the fraction of the globe that is covered each year by in situ sampling and by the magnitude of the mean shift (B7). The result may not always be intuitive. For example, in 2010, the 0–1800-m Clim1950 ZIF value is 50 ZJ less than the ClimArgo ZIF value (Fig. 4, panel d, dashed lines), despite Argo’s near global coverage in 2010. We can compute the expected size of the shift from B10: The mean shift that the Clim1950 curve experiences in 2010 is  $(1 - f_{0-1800}^{2010})m^{1950,0-1800}A$ . In this case,  $m^{1950,0-1800}A \approx -500$  ZJ (Fig. 4, panel d) and  $f_{0-1800}^{2010} \approx .9$  (Fig. 3), so the observed offset in 2010 of 50 ZJ is to be expected. An intermediate climatology, say one shifted to 1980 levels, would result in a smaller cold bias in modern times, but a large warm bias in earlier times.

Unlike the ZIF estimates, the REP estimates (Fig. 4, black solid line) are the same for both the Clim1950 values and the ClimArgo (B11). This result holds because the REP estimates avoid any relaxation back to climatological values by assuming that the mean anomaly of the sampled regions applies to the unsampled regions.

## 6. Sampling uncertainty

Computing the global integral using the REP estimates increases both the signal within and the noise from the observed areas of the global ocean by extending both quantities to the unmeasured areas (Lyman and Johnson 2008). The sampling uncertainty (Fig. 5),

which generally grows as sampling becomes sparse (Fig. 3), is only one component of the total uncertainty (e.g., Lyman et al. 2010). Here we do not concern ourselves with other contributions, but there is much work being done by the research community on quantifying uncertainties remaining in XBT and MBT bias corrections (see Cowley et al. 2013 for a recent example), as well as those owing to differences in quality control, mapping, and climatology.

Sampling uncertainty (Fig. 5) increases greatly with decreased area coverage (Fig. 3), so it ramps up going back in time prior to about 1956 in the 0–100-m layer. After that date, around the time of the International Geophysical Year (IGY), MBT use expanded. The advent of the IGY further reduces sampling uncertainties in all layers because of an increase in reversing thermometer data from deep oceanographic stations. The sampling uncertainty for 0–300 m declines further after around 1969, with the increasingly widespread use of the shallow XBT, and decreases markedly for 0–700 m around 1990, with the more widespread use of the deep XBT and the advent of WOCE. Again, WOCE reduces sampling uncertainty in all layers during the 1990s because of increased CTD stations worldwide. However, around the year 2000, there is a slight increase in sampling uncertainty in all layers as WOCE sampling winds down before Argo sampling spins up. Starting around 2001–2003, sampling uncertainties in all layers begin to shrink, reaching a historical low by 2005–2006, depending on the depth of integration, and remaining small as Argo coverage continues to improve.

Sampling uncertainties (Fig. 5) generally decrease as areal coverage (Fig. 3) increases. This inverse relationship is not perfect, at least partly because OHCA variance is spatially variable, so increasing sampling in one region might reduce sampling

uncertainty more than increasing it the same amount in another region. Furthermore, the contribution to sampling uncertainty in deeper layers is relatively small even though the area sampled is also small, except in recent years. This situation likely holds owing to the decreasing heat content anomaly signals of eddies, planetary waves, gyre shifts, and even the net warming signal with increasing depth.

## 7. Discussion

Ocean warming is observed between 1950 and 2011 in all layers for both methods of global integration: zero-infilled (ZIF) or representative (REP) means (Fig. 4). However, the REP estimates are more likely to result in an unbiased trend (Lyman and Johnson 2008). Furthermore, we show that the REP estimates are insensitive to mean shifts in the background climatology, whereas the trends from the ZIF estimates vary greatly depending on such shifts. For a modern (warmer) climatology, ZIF estimates for early, poorly sampled years shift back toward zero anomaly (warm conditions), increasingly reducing the warming trend the further back in time that trend is computed. In contrast, for climatologies shifted to early (colder) values, ZIF estimates for early years are cold, but more appropriately warm in later years because the data coverage of the warmed oceans is improved in recent years. This difference results in a very different trend depending on the climatology for ZIF estimates. While further analyses are needed to quantify the impacts of regional variability in climatologies (rather than the global shifts applied here) on the OHCA curves, REP estimates provide an estimate of the warming rate that does not depend on the mean heat content of the climatology. To the extent that the oceans are warming more-or-less uniformly within depth layers, REP estimates are preferable to ZIF estimates.

Here we only compute and analyze sampling uncertainties. Before the significance of the trends in the global integral OHCA can be completely determined over various lengths of the historical record, it will be necessary to examine other sources of uncertainty in more detail. These uncertainties include factors owing to quality control, choice of XBT and MBT bias correction, choice of climatology, and choice of mapping method. Nonetheless, when areal coverage gets smaller, OHCA sampling uncertainties generally get larger. There is a roughly inverse relationship between these quantities in each layer, with smaller and smaller contributions of the uncertainty per unit depth with increasing depth. Of course, the warming signal in the upper 1800 m also gets smaller per unit depth with increasing depth (e.g., Johnson and Wijffels 2011). The 50% coverage benchmark (Fig. 4, vertical lines) discussed above for the different layers is arbitrary, but a reminder that below some areal coverage, it becomes difficult to make a realistic estimate of the global integral OHCA for a given year in a given layer, no matter the mapping method.

In Lyman and Johnson (2008), a similarly arbitrary benchmark of 10 ZJ was used to highlight that sampling before 1967 produced large uncertainties. That benchmark (Fig. 1 of Lyman and Johnson 2008) roughly corresponds to the benchmark used here of 50% coverage in the deepest layer under consideration. For this analysis increases in coverage for the deepest layer under consideration correspond to decreases in sampling uncertainty from the surface down to that layer (Fig. 6). The mean of the 0–700-m layer sampling uncertainty between 45% and 55% coverage for the 300–700-m layer is  $11.4 \pm 1.4$  ZJ, similar to the Lyman and Johnson (2008) benchmark.

Note that in Lyman and Johnson (2008), shallow XBTs were extended to 700 m following Willis et al. (2004), but that extrapolation technique is not employed here. As a result, before widespread use of the deep XBT, the 0–700-m sampling density in Lyman and Johnson (2008) roughly corresponds to the 100–300-m sampling density here, which reaches 50% in 1967. The 300–700-m sampling density here does not reach 50% until 1983, when the deep XBT came into widespread use.

Computation of the sampling uncertainty requires the assumption that SSH is directly proportional to OHCA in a depth layer. While this is a reasonable assumption, the two are not directly related because SSH also depends on OHCA over the entire water column, changes in salinity, and mass changes. On the other hand, sampling density is a product of the mapping and requires no assumption about the relationship of SSH to OHCA. Finally, it would seem unwise to apply infilling assumptions to over 50% of the mapping area when computing a global integral; thus, the 50% benchmark is an arbitrary, but perhaps reasonable, choice.

We estimate warming rates for the different layers by computing trends using unweighted least squares REP OHCA annual values from the benchmark years of 1956, 1967, 1983, and 2004, through the final year, 2011 (Table 1). All warming estimates presented in this paper in  $W\ m^{-2}$  are calculated as means for the Earth's entire surface area. We do not use a weighted least squares formalism or report uncertainties because the sampling uncertainties estimated here are not the full uncertainty, which can be much larger (Lyman et al. 2010). For adequately sampled time periods, warming trends generally increase with increasing depth, as the surface-intensified warming signal penetrates to at least 1500 m (e.g., Levitus et al. 2012). In recent years, from 2004 to

2011, while the upper ocean is not warming, the ocean continues to absorb heat at depth (e.g., Levitus et al. 2012; von Schuckman and Le Traon 2011), here estimated at a rate of  $0.56 \text{ W m}^{-2}$  when integrating over 0–1800 m. The warming estimated here for 0–700-m REP from 1993 through 2011 is  $0.56 \text{ W m}^{-2}$ , whereas that for our previously published curve is  $0.48 \text{ W m}^{-2}$  (Johnson et al. 2012). The historical rates for 0–100, 0–300, and 0–700 m estimated here are generally larger than other analyses (e.g., Cowley et al. 2013; Levitus et al. 2012; Ishii and Kimoto 2009; Domingues et al. 2008), at least partly owing to our use of REP, which increases the trend by assuming the anomaly in unsampled areas is the same as the mean anomaly of sampled areas. Our ZIF estimates of the warming from 1955 to 2010 range from  $0.05\text{--}0.29 \text{ W m}^{-2}$  for 0–700 m and  $0.05\text{--}0.34 \text{ W m}^{-2}$  for 0–1800 m (Fig. 4). This range is consistent with Levitus et al. (2012), which reports warming rates of  $0.19 \text{ W m}^{-2}$  for 0–700 m and  $0.27 \text{ W m}^{-2}$  for 0–2000 m over the same period.

While XBT biases have received much attention in the last few years, MBTs are a large source of upper ocean temperature data from the 1950s through the 1960s, and also appear to have large, possibly time-dependent biases (e.g., Gouretski and Reseghetti 2010; Hamon et. al 2012; Ishii and Kimoto, 2009; Levitus et. al 2009; Levitus et. al 2012). Sampling uncertainties for the REP estimate during this early time period are large (Fig. 5); however, the warming signal is also large (Fig. 4). In order to determine if this large signal is statistically significant, it will be necessary to estimate all possible sources of uncertainty, including uncertainties in MBT bias corrections.

Finally, the results presented here hold generally for any linear mapping (Appendix B) but the magnitudes of the results are applicable only to the particular

mapping scheme used here. Changing the mapping scheme, including using different time or length scales or different mapping formalism will alter the results quantitatively, but not alter the conclusions qualitatively.

*Acknowledgments.*

Altimeter products used herein were produced by SSALTO/DUACS as part of the Environment and Climate EU Enact project (EVK2–CT2001–00117) and distributed by AVISO, with support from CNES. The bulk of the in situ data used herein was provided by Hadley Centre ([www.metoffice.gov.uk/hadobs](http://www.metoffice.gov.uk/hadobs)). Float data were collected and made freely available by the Argo Program (of the Global Ocean Observing System) and contributing national programs (<http://www.argo.net/>). Comments of three anonymous reviewers improved the manuscript. The NOAA Climate Program Office and the NOAA Office of Oceanic and Atmospheric Research provided support for this research. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the National Oceanic and Atmospheric Administration.

## 505 APPENDIX A. Summation of Uncertainty

### 506 *Sampling uncertainty computation for a general linear mapping*

507 In using Aviso SSH as a proxy for every depth layer of OHCA (following Lyman and  
 508 Johnson 2008), we implicitly assume that eddy statistics of the upper layers are similar to  
 509 deep layers. In each depth layer the local correlations of Aviso SSH with OHCA are  
 510 estimated on a coarse grid. The resulting coefficients are used to construct proxy OHCA  
 511 fields that are then subsampled in locations of in situ observation. Sampling uncertainty  
 512 in a depth layer ( $SE_j$ , where the integer numbers 1, 2, 3, 4, and 5 have been assigned the  
 513 0–100, 100–300, 300–700, 700–900, and 900–1800-m layers, respectively), is defined as  
 514 the standard deviations of the differences between heat content curves computed from the  
 515 proxy OHCA fields and the in situ OHCA fields (Lyman and Johnson 2008).

### 516 *Addition of sampling uncertainty when summing depth layer OHCA*

517 We also make another assumption, that the correlated sampling uncertainty in depth layer  
 518  $j$  between depth layers  $i$  and  $j$  ( $CorrSE_{ij}$ ) is proportional to the square root of the

519 percentage change in the globe coverage for depth layer  $i$  ( $\frac{\Delta_i^y}{f_j^y}$ ):

$$520 \quad CorrSE_{ij} \equiv SE_j \sqrt{\frac{\Delta_i^y}{f_j^y}}, \quad i \geq j, \quad (A1)$$

$$0, \quad i < j$$

521

522 where  $\Delta_i^y \equiv f_i^y - f_{i-1}^y$ . Note that  $\Delta_5^y = f_5^y$ , because layer 5 is the bottom layer and

$$523 \quad \sum_{l=1}^5 \Delta_l^y = f_1^y.$$



524 The fractional coverage ( $f_l^y$ ) for depth layer  $l$  and year  $y$  is defined as

$$525 \quad f_l^y = \frac{\sum_{i=0, I}^{j=0, J} \langle 1_k \rangle_{i,j} dA_{i,j}}{A}, \quad (A2)$$

526 where  $\langle 1_k \rangle_{i,j}$  is the sampling density for a given data distribution,  $obs_k$ . The sampling  
 527 density is a linear mapping to location  $i, j$  where each datum at position  $k$ , for depth  
 528 layer  $l$  and year  $y$ , has been replaced by unity. Here  $dA_{i,j}$  is the area represented by the  
 529 location  $i, j$ , and  $A \equiv \sum_{i=0, I}^{j=0, J} dA_{i,j}$ , the total area mapped (taken from Lyman and Johnson  
 530 2008). Also,  $f_l^y$  and  $\langle 1_k \rangle_{i,j}$  quantify the density of the sampling based on the data  
 531 distribution and the particulars of temporal and spatial scales of whatever mapping  
 532 scheme is being employed (Lyman and Johnson 2008). This methodology is a more  
 533 quantitative assessment of sampling density than other methods sometimes used, because  
 534 it depends on the particulars of the mapping scheme (including correlation functions,  
 535 covariance length-scales, and signal-to-noise variance ratios for a standard objective map  
 536 – e.g., Bretherton et al. 1976) rather than, for example, some arbitrary threshold such as a  
 537 floor on the number of data points contained in a box for that box to be included in the  
 538 calculation.

539 Summing the correlated sampling uncertainties from the surface to the bottom of  
 540 a given depth layer and then adding the resulting uncertainties in quadrature (thus  
 541 assuming they are uncorrelated) yields:

$$542 \quad TSE_m = \sqrt{\sum_{i=1}^5 \left[ \sum_{j=1}^m CorrSE_{ij} \right]^2}, \quad (A3)$$

543    where  $TSE_m$  is the total sampling uncertainty from the surface to the bottom of layer  $m$ .

544

## APPENDIX B. Global Integrals

*General solution for the effects of uniform climatological shifts on global integrals*

Two global integrals are examined in this paper, one based on the zero-infilled integrals or means (*ZIF*) and the other based on the representative integrals or means (*REP*). They can be defined as

$$ZIF = \sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k \rangle_{i,j} dA_{i,j} \quad (B1)$$

and

$$REP = \frac{A \sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k \rangle_{i,j} dA_{i,j}}{\sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j}}, \quad (B2)$$

where  $\langle obs_k \rangle_{i,j}$  is any linear mapping of some distribution of observation ( $obs_k$ ) on a depth layer  $l$  and for year  $y$  to a location  $i, j$  and the other terms are defined in Appendix A.

If the  $obs_k$  on depth layer  $l$  are shifted by a constant value, in this case the representative mean for year  $ym$ , then

$$m_k^{ym,l} \equiv \frac{\sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k \rangle_{i,j} dA_{i,j}}{\sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j}}. \quad (B3).$$

Here the mapping  $\langle obs_k \rangle_{i,j}$  is dependent on year,  $ym$ , and depth layer,  $l$ . Adding this shift

B1 and B2 become

$$ZIF^s = \sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k - m_k^{ym,l} \rangle_{i,j} dA_{i,j} \quad (B4)$$

and

$$REP^s = \frac{A \sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k - m_k^{ym,l} \rangle_{i,j} dA_{i,j}}{\sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j}}, \quad (B5)$$

where  $ZIF^s$  and  $REP^s$  are integrals that have been shifted by this spatially uniform mean. Since the mappings are linear and  $m_k^{ym,l}$  is spatially constant, therefore not dependent on  $k$ , this term can be moved outside the sum so that B4 and B5 become

$$ZIF^s = \sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k \rangle_{i,j} dA_{i,j} - m_k^{ym,l} \sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j} \quad (B6)$$

and

$$REP^s = \frac{A \sum_{\substack{i=0,I \\ j=0,J}} \langle obs_k \rangle_{i,j} dA_{i,j}}{\sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j}} - \frac{Am_k^{ym,l} \sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j}}{\sum_{\substack{i=0,I \\ j=0,J}} \langle 1_k \rangle_{i,j} dA_{i,j}}. \quad (B7)$$

Eq. B6 and B7 simplify to

$$ZIF^s = ZIF - f_l^y m^{ym,l} A \quad (B8)$$

and

$$REP^s = REP - m^{ym,l} A, \quad (B9)$$

where  $f_l^y$  is the fractional coverage (A2). Thus  $REP^s$  has shifted by area integral of representative mean for year  $ym$ ,  $m^{ym,l} A$  (B2 evaluated at  $ym$ ). This shift is independent of the sampling in the mapping and therefore independent of the year that of the map,  $y$ .

577 On the other hand,  $ZIF^s$  is shifted by the fraction of the area integral of representative  
 578 mean that was sampled,  $f_i^y m^{ym,l} A$ , and is dependent on  $y$ . These results are independent  
 579 of mapping or mean offset.

580 When plotting shifted OHCA curves, it is useful to offset the results in order to  
 581 focus attention on changes in trends rather than simple offsets. Here we have chosen to  
 582 offset the shifted OHCA curves by the area integral of the climatology offset,  $m^{ym,l} A$ . For  
 583 plotting purposes, B8 and B9 now become:

$$584 \quad ZIF^s = ZIF + m^{ym,l} A (1 - f_i^y) \quad (B10)$$

585 and

$$586 \quad REP^s = REP. \quad (B11)$$

587 Thus, the  $REP$  result is independent of baseline shifts in climatology, whereas the  $ZIF$   
 588 result for any given year is sensitive to baseline shifts proportional to the unsampled area  
 589 for that year.

590

## References

- Barker, P. M., J. R. Dunn, C. M. Domingues, and S. E. Wijffels, 2011: Pressure sensor drifts in Argo and their impacts. *J. Atmos. Oceanic Technol.*, **28**, 1036–1049, doi:10.1175/2011JTECHO831.1.
- Bindoff, N. L., and Coauthors, 2007: Observations: Oceanic climate change and sea level. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 385–432.
- Boyer, T. P., and Coauthors, 2009: *World Ocean Database 2009*. S. Levitus, Ed., NOAA Atlas NESDIS 66, U.S. Gov. Printing Office, Wash., D.C., 216 pp., DVDs.
- Bretherton, F. P., R. E. Davis, and C. B. Fandry, 1976: A technique for objective analysis and design of oceanographic experiment applied to MODE-73. *Deep-Sea Research*, **23**, 559–582.
- Cowley, R., S. Wijffels, L. Cheng, T. Boyer, and S. Kizu, 2013: Biases in Expendable BathyThermograph data: a new view based on historical side-by side comparisons. *J. Atmos. Oceanic Technol.*, **30**, 1195–1225, doi:10.1175/JTECH-D-12-00127.1.
- Carton, J. A., B. S. Giese, X. Cao, and L. Miller, 1996: Impact of altimeter, thermistor, and expendable bathythermograph data on retrospective analyses of the tropical Pacific Ocean. *J. Geophys. Res.*, **101**, 14 147–14 159, doi:10.1029/96JC00631.
- Domingues, C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker, and J. R. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea level rise. *Nature*, **453**, 1090–1094, doi:10.1038/nature07080.

- 612 Emery, W. J., and R. E. Thomson, 1998: *Data Analysis Methods in Physical*  
 613 *Oceanography*. Elsevier, Oxford, 634 pp.
- 614 Gleckler, P.J., B. D. Santer, C. M. Domingues, D.W. Pierce, T. P. Barnett, J. A. Church,  
 615 K. E. Taylor, K. M. AchutaRao, T. P. Boyer, M. Ishii and P. M. Caldwell, 2012:  
 616 Human-induced global ocean warming on multidecadal timescales. *Nature Climate*  
 617 *Change*, **2**, 524–529, doi: 10.1038/NCLIMATE1553.
- 618 Gouretski, V., and F. Reseghetti, 2010: On depth and temperature biases in  
 619 bathythermograph data: Development of a new correction scheme based on analysis  
 620 of a global ocean database. *Deep-Sea Res. I*, **57**, 812–833,  
 621 doi:10.1016/j.dsr.2010.03.011.
- 622 Gouretski, V., J. Kennedy, T. Boyer, and A. Köhl, 2012: Consistent near-surface ocean  
 623 warming since 1900 in two largely independent observing networks. *Geophys. Res.*  
 624 *Lett.*, **39**, L19606, doi:10.1029/2012GL052975.
- 625 Hamon, M., G. Reverdin, and P.-Y. Le Traon, 2012: Empirical Correction of XBT Data.  
 626 *J. Atmos. Oceanic Technol.*, **29**, 960–973, doi: 10.1175/JTECH-D-11-00129.1.
- 627 Harrison, D. E., and N. K. Larkin, 1998: El Niño-Southern Oscillation sea surface  
 628 temperature and wind anomalies, 1946–1993. *Rev. Geophys.*, **36**, 353–399,  
 629 doi:10.1029/98RG00715.
- 630 Ingleby, B., and M. Huddleston, 2007: Quality control of ocean temperature and salinity  
 631 profiles—historical and real-time data. *J. Mar. Syst.*, **65**, 158–175, doi:  
 632 10.1016/j.jmarsys.2005.11.019.

- 633 Ishii, M., and M. Kimoto, 2009: Reevaluation of historical ocean heat content variations  
 634 with time-varying XBT and MBT depth bias corrections. *J. Oceanogr.*, **65**, 287–299,  
 635 doi:10.1007/s10872-009-0027-7.
- 636 Johnson, G. C., J. M. Lyman, J. K. Willis, S. Levitus, T. Boyer, J. Antonov, and S. A.  
 637 Good, 2012: Global Oceans: Ocean Heat Content. In State of the Climate in 2011,  
 638 Blunden, J., and D. S. Arndt, Eds., *Bulletin of the American Meteorological Society*,  
 639 **93**, 7, S62-S65, doi:10.1175/2012BAMSStateoftheClimate.1.
- 640 Johnson, G. C., M. J. McPhaden, G. D. Rowe, and K. E. McTaggart, 2000: Upper  
 641 equatorial Pacific Ocean current and salinity variability during the 1996–1998 El  
 642 Niño–La Niña cycle. *J. Geophys. Res.*, **105**, 1037–1053, doi:10.1029/1999JC900280.
- 643 Johnson, G. C., and S. E. Wijffels, 2011: Ocean density change contributions to sea level  
 644 rise. *Oceanogr.*, **24**(2), 112–121, doi:10.5670/oceanog.2011.31.
- 645 Kouketsu, S., and Coauthors, 2011: Deep ocean heat content changes estimated from  
 646 observation and reanalysis product and their influence on sea level change. *J.*  
 647 *Geophys. Res.*, **116**, C03012, doi:10.1029/2010JC006464.
- 648 Larkin, N. K., and D. E. Harrison, 2001: Tropical Pacific ENSO cold events, 1946–95:  
 649 SST, SLP, and surface wind composite anomalies, *J. Climate*, **14**, 3904–3931,  
 650 doi:10.1175/1520-0442(2001)014<3904:TPECES>2.0.CO;2.
- 651 Levitus, S., J. Antonov, and T. Boyer, 2005: Warming of the World Ocean, 1955–2003.  
 652 *Geophys. Res. Lett.*, **32**, L02604, doi:10.1029/2004GL021592.
- 653 Levitus, S., J. I. Antonov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, and A. V.  
 654 Mishonov, 2009: Global ocean heat content 1955–2008 in light of recently revealed  
 655 instrumentation problems. *Geophys. Res. Lett.*, **36**, L07608,



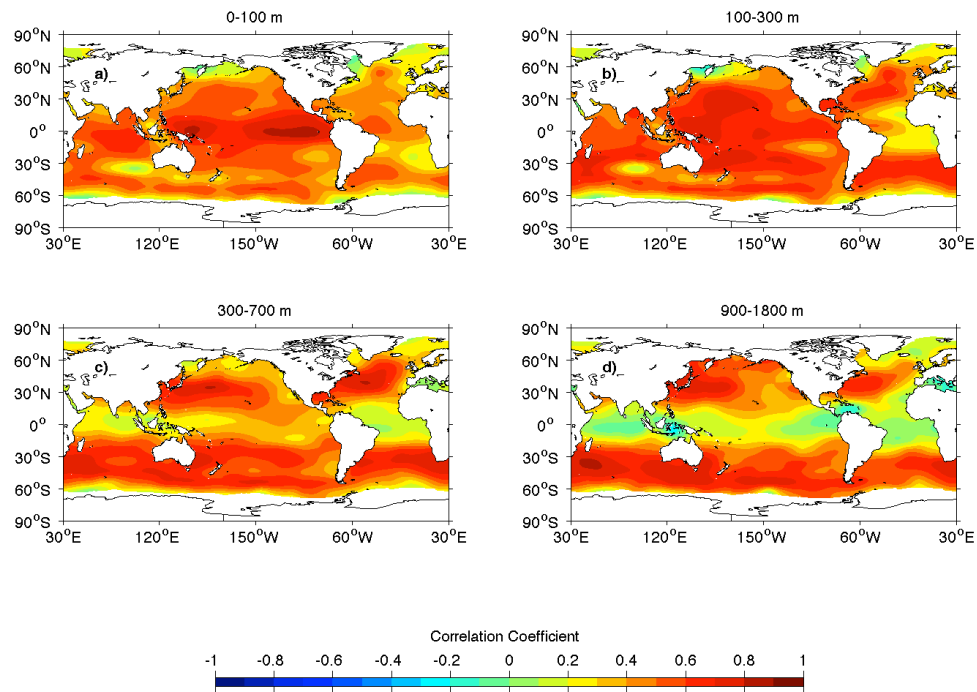
- doi:10.1029/2008GL037155.
- Levitus, S., and Coauthors, 2012: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, L10603, doi:10.1029/2012GL051106.
- Llovel, A., I. Funkumori, and B. Meyssignax, 2013: Depth-dependent temperature change contributions to global mean thermosteric sea level rise from 1960 to 2010. *Global and Planetary Change*, **101**, 113–118, doi:/10.1016/j.gloplacha.2012.12.011.
- Lozier, M. S., S. J. Leadbetter, R. G. Williams, V. Roussenov, M. S. C. Reed, and N. J. Moore, 2008: The spatial pattern and mechanisms of heat content change in the North Atlantic. *Science*, **319**, 800–803, doi:10.1126/science.1146436.
- Lyman, J. M., and G. C. Johnson, 2008: Estimating annual global upper ocean heat content anomalies despite irregular in situ ocean sampling. *J. Climate*, **21**, 5629–5641, doi:10.1175/2008JCLI2259.1.
- Lyman, J. M., S. A. Good, V. V. Gouretski, M. Ishii, G. C. Johnson, M. D. Palmer, D. M. Smith, and J. K. Willis, 2010: Robust warming of the global upper ocean. *Nature*, **465**, 334–337, doi:10.1038/nature09043.
- Mauritzen, C., A. Melsom, and R. T. Sutton, 2012: Importance of density-compensated temperature change for deep North Atlantic Ocean heat uptake, *Nature Geosci.*, **5**, 905–910, doi:10.1038/ngeo1639.
- Palmer, M. D., and P. Brohan, 2011: Estimating sampling uncertainty in fixed-depth and fixed-isotherm estimates of ocean warming. *Int. J. Climatol.* **31**, 980–986, doi:10.1002/joc.2224.

- 678 Palmer, M. D., K. Haines, S. F. B. Tett, and T. J. Hansell, 2007: Isolating the signal of  
 679 ocean global warming. *Geophys. Res. Lett.*, **34**, L23610, doi:  
 680 10.1029/2007GL031712.
- 681 Palmer, M. D., and Coauthors, 2010: Future observations for monitoring global ocean  
 682 heat content, in *Proceedings of the OceanObs' 09: Sustained Ocean Observations and*  
 683 *Information for Society Conference* (Vol. 2), Venice, Italy, 21–25 September 2009, J.  
 684 Hall, D. E. Harrison, and D. Stammer, Eds., ESA Publication WPP-306.
- 685 Purkey, S. G., and G. C. Johnson, 2010: Warming of global abyssal and deep Southern  
 686 Ocean between the 1990s and the 2000s: Contributions to global heat and sea level  
 687 rise budgets. *J. Climate*, **23**, 6336–6351, doi:10.1175/2010JCLI3682.1.
- 688 Roemmich, D., and J. Gilson, 2009: The 2004–2008 mean and annual cycle of  
 689 temperature, salinity, and steric height in the global ocean from the Argo Program  
 690 *Prog. Oceanogr.*, **82**, 81–100, doi:10.1016/j.pocean.2009.03.004.
- 691 Roemmich, D., and J. Gilson, 2011: The global ocean imprint of ENSO, *Geophys. Res.*  
 692 *Lett.*, **38**, L13606, doi: 10.1029/2011GL047992.
- 693 Roemmich, D., and Coauthors, 2009: The Argo Program: Observing the global oceans  
 694 with profiling floats. *Oceanogr.*, **22(2)**, 34–43, doi:10.5670/oceanogr.2009.36.
- 695 Roemmich, D., W. J. Gould, and J. Gilson, 2012: 135 years of global ocean warming  
 696 between the Challenger expedition and the Argo Programme. *Nature Climate*  
 697 *Change*, **2**, 425–428, doi:10.1038/nclimate1461.
- 698 Stuecker, M. F., A. Timmermann, F.-F. Jin, S. McGregor, and H.-L. Ren, 2013: A  
 699 combination mode of the annual cycle and the El Niño/Southern Oscillation. *Nature*  
 700 *Geosci.*, doi:10.1038/ngeo1826.

- 701 von Schuckmann, K., and P.-Y. Le Traon, 2011: How well can we derive global ocean  
702 indicators from Argo data? *Ocean Sci.*, **7**, 783–791, doi:10.5194/os-7-783-2011.
- 703 von Schuckmann, K., F. Gaillard, and P.-Y. L. Traon, 2009: Global hydrographic  
704 variability patterns during 2003–2008. *J. Geophys. Res.*, **114**, C09007,  
705 doi:10.1029/2008JC005237.
- 706 Warren, B. A., 2008: Nansen-bottle stations at the Woods Hole Oceanographic  
707 Institution. *Deep-Sea Res. I*, **55**, 379–395, doi: /10.1016/j.dsr.2007.10.003.
- 708 Wijffels, S. E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K.  
709 Ridgway, and J. A. Church, 2008: Changing Expendable Bathythermograph Fall  
710 Rates and Their Impact on Estimates of Thermosteric Sea Level Rise, *J. Climate*, **21**,  
711 5657–5672, doi: 10.1175/2008JCLI2290.1.
- 712 Willis, J. K., D. Roemmich, and B. Cornuelle, 2004: Interannual variability in upper  
713 ocean heat content, temperature, and thermosteric expansion on global scales, *J.*  
714 *Geophys. Res.*, **109**, C12036, doi:10.1029/2003JC002260.
- 715

Table 1. Warming reported as heat flux applied to Earth's entire surface area (in  $\text{W m}^{-2}$ ) corresponding to trends in annual REP OHCA estimates from unweighted linear fits from benchmark years through 2011 for different depths of integration (left column). A benchmark year is defined as the year in which sampling coverage all layers being considered first exceeds 50%, and remains  $> 50\%$  thereafter. Layer warming trends over time periods during which coverage in a layer is  $< 50\%$  in any year, indicated here by a dash, are not reported.

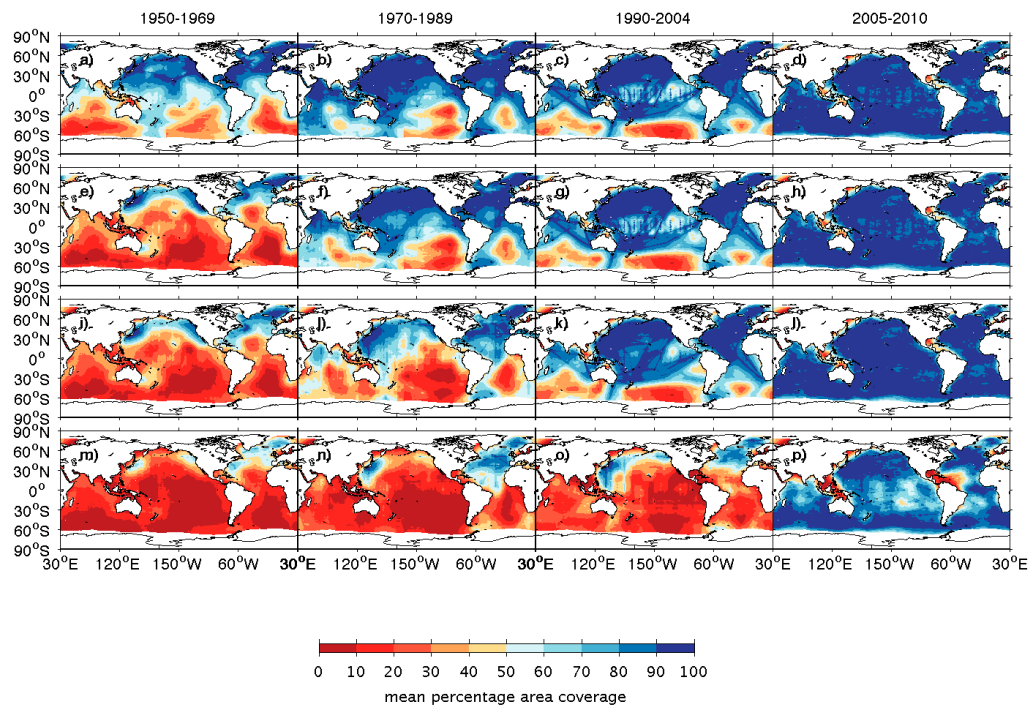
Depth layer	Time Period			
	1956–2011	1967–2011	1983–2011	2004–2011
0–100 m	$0.06 \text{ W m}^{-2}$	$0.08 \text{ W m}^{-2}$	$0.08 \text{ W m}^{-2}$	$0.01 \text{ W m}^{-2}$
0–300 m	–	$0.18 \text{ W m}^{-2}$	$0.24 \text{ W m}^{-2}$	$0.19 \text{ W m}^{-2}$
0–700 m	–	–	$0.46 \text{ W m}^{-2}$	$0.30 \text{ W m}^{-2}$
0–1800 m	–	–	–	$0.56 \text{ W m}^{-2}$



723

724 FIG. 1. Correlation coefficients between Aviso SSH and (a) 0–100-m OHCA, (b) 100–  
 725 300-m OHCA, (c) 300–700-m OHCA, and (d) 900–1800-m OHCA. Correlation  
 726 coefficients are computed using the same methods as for regression coefficients (Section  
 727 3c).

728



729

730 FIG. 2. Means of annual in situ temperature data sampling density (see A3) for various  
 731 depth layers and time ranges. From top to bottom: 0–100 m, 100–300 m, 300–700 m, and  
 732 900–1800 m. From left to right: 1950–1969, 1970–1989, 1990–2004, and 2005–2010.

733

734

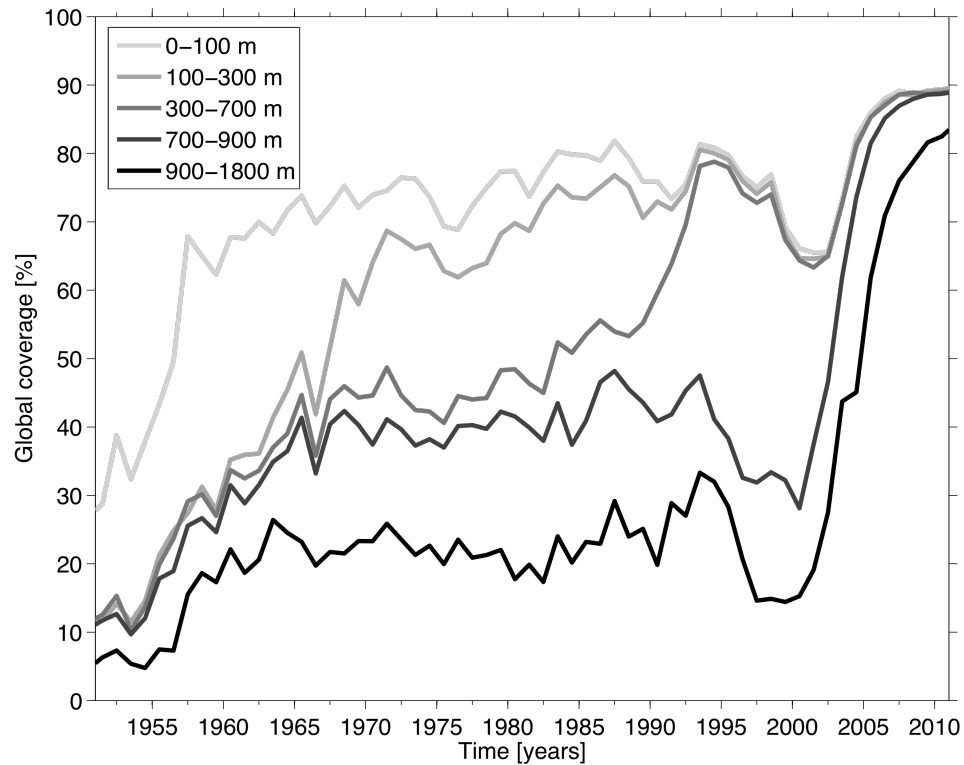


FIG. 3. Annual fractional coverage of global ice-free ocean sampled for in situ ocean temperature (see A2) for (proceeding from lightest grey to black lines) 0–100 m, 100–300 m, 300–700 m, 700–900 m, and 900–1800 m.

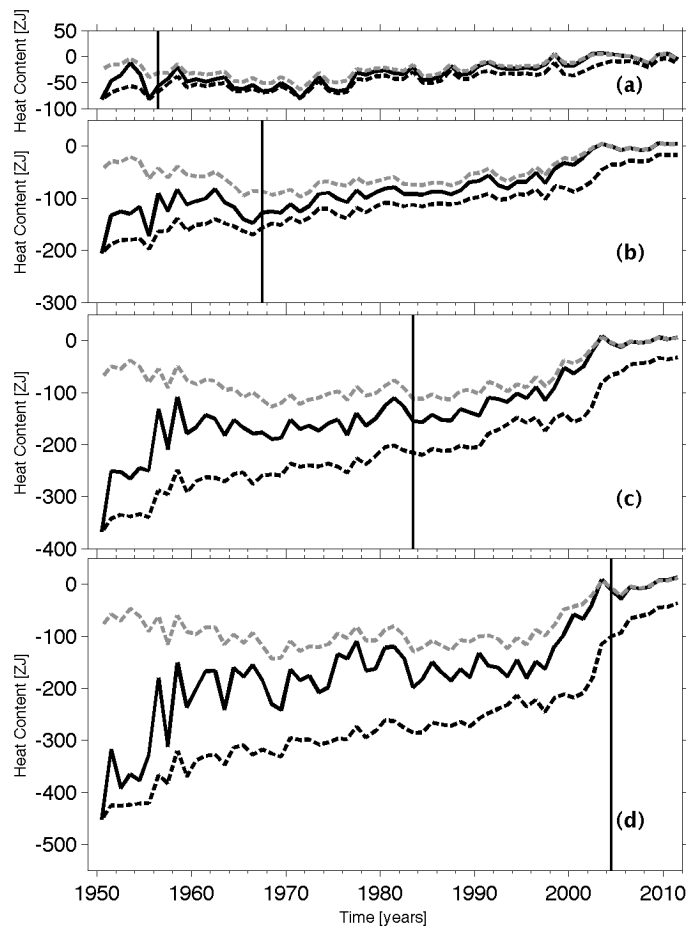
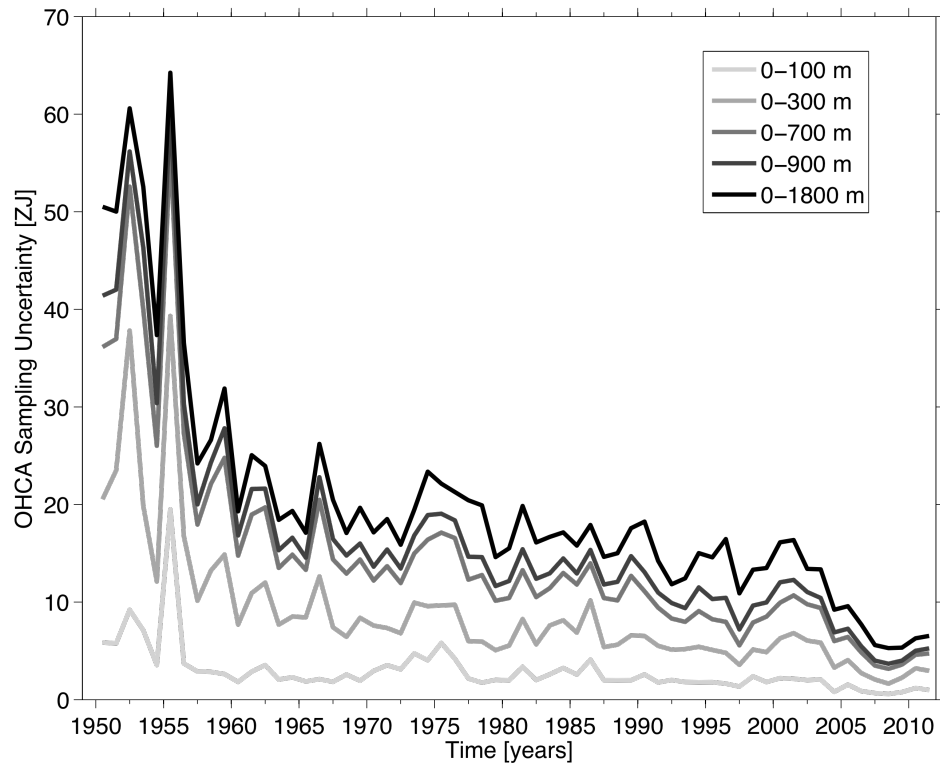


FIG. 4. Time series of annual average global integrals of upper ocean heat content anomaly ( $10^{21}$  J, or ZJ) for (a) 0–100 m, (b) 0–300 m, (c) 0–700 m, and (d) 0–1800 m. Time series are shown using ZIF estimates relative to both ClimArgo (dashed grey lines) and Clim1950 (dashed black lines). Time series are also shown using REP estimates (black solid lines), which are not affected by shifts in the mean climatology (B11). Thin vertical lines denote when the coverage (Fig. 3) reaches 50% for (a) 0–100 m, (b) 100–300 m, (c) 300–700 m, and (d) 900–1800 m.



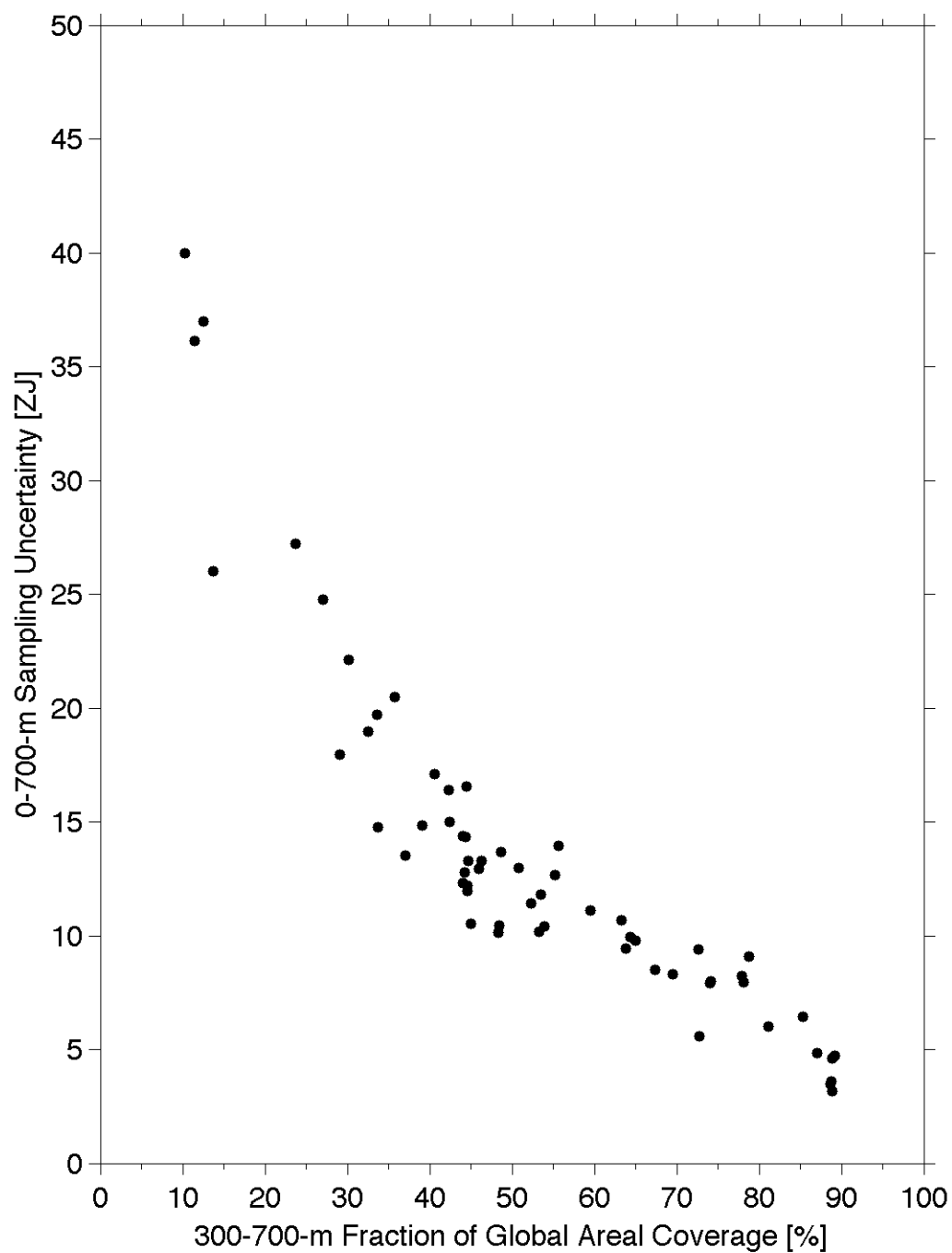
749



750

751 FIG. 5. Estimates of sampling uncertainties for the ZIF estimates (following Lyman and  
 752 Johnson 2008) at one standard error of the mean for (proceeding from lightest grey to  
 753 black lines) 0–100 m, 0–300 m, 0–700 m, 0–900 m, and 0–1800 m. Uncertainties are  
 754 computed by summing up sampling uncertainties for each layer from the surface  
 755 downwards (Appendix A). Where sampled layers overlap in space, uncertainties are  
 756 added. Where sampled layers do not spatially overlap, uncertainties are summed in  
 757 quadrature.

758



758

759 FIG. 6. 0–700-m sampling uncertainty (Fig. 5) plotted vs. 300–700-m annual fractional

760 coverage (Fig. 3).